LINEAR ALGEBRA AND ITS APPLICATIONS 5, 19–28 (1972)

An Inequality for Schur Functions

MARVIN MARCUS AND HENRYK MINC*
University of California
Santa Barbara, California

ABSTRACT

If H is a subgroup of the symmetric group of degree n and χ is a complex character on H of degree 1, then the Schur function for H and χ is defined by

$$d_\chi^H(Y) = \sum_{\sigma\in H} \chi(\sigma) \prod_{i=1}^{n} y_{i\sigma(i)}$$

for any n-square matrix $Y = (y_{ij})$.

It is shown that, if A_1 is a positive definite matrix, A_2 a positive semidefinite nonzero matrix, and μ_1, μ_2 complex numbers, then

$$|d_\chi^H(\mu_1 A_1 + \mu_2 A_2)| \leqslant d_\chi^H(|\mu_1| A_1 + |\mu_2| A_2).$$

Other inequalities relating functions of $\mu_1 A_1 + \mu_2 A_2$ and of $|\mu_1| A_1 + |\mu_2| A_2$ are also obtained.

1. STATEMENTS OF RESULTS

Let H be a subgroup of order h of S_n, the symmetric group of degree n, and let χ be a character of degree 1 on H. Then, for $Y = (y_{ij})$, an n-square matrix, let

$$d_\chi^H(Y) = \sum_{\sigma\in H} \chi(\sigma) \prod_{i=1}^{n} y_{i\sigma(i)}, \tag{1}$$

the generalized matrix function of Schur [**3**].

* The work of both authors was supported by the U.S. Air Force Office of Scientific Research under Grant AFOSR 698-67.

In what follows, μ_1 and μ_2 are nonzero complex numbers, A_1 is an n-square positive definite hermitian matrix, and A_2 is an n-square positive semidefinite hermitian matrix. We indicate this with the notation $A_1 > 0$, $A_2 \geqslant 0$. In general, $X \geqslant Y$ will mean that $X - Y$ is positive semidefinite hermitian.

Our first result follows.

THEOREM 1. *If $A_1 > 0$ and $A_2 \geqslant 0$ $(A_2 \neq 0)$, then*

$$|d_\chi^H(\mu_1 A_1 + \mu_2 A_2)| \leqslant d_\chi^H(|\mu_1| A_1 + |\mu_2| A_2). \tag{2}$$

If $A_2 > 0$, then equality holds in (2) if and only if $\mu_1/\mu_2 > 0$. If $\chi \equiv 1$, the condition $\mu_1/\mu_2 > 0$ is necessary and sufficient for equality in (2), even if A_2 is singular.

In the case that $H = S$, $\chi = \operatorname{sgn}$, $d_\chi^H = \det$, the inequality (2) is due to W. M. Frank [1].

In order to state our next result concerning S-functions, we require some combinatorial notation. For $1 \leqslant m \leqslant n$, let $\Gamma_{m,n}$ be the totality of sequences $\omega = (\omega(1), \ldots, \omega(m))$, $1 \leqslant \omega(i) \leqslant n$, $i = 1, \ldots, m$. We say that $\omega \sim \tau$ if and only if there exists $\sigma \in H$ such that $\omega\sigma = (\omega(\sigma(1)), \ldots, \omega(\sigma(m))) = \tau$. For fixed m, n, and H we denote by Δ a system of distinct representatives for $\sim$, so chosen that each element of Δ is first lexicographically in its equivalence class in $\Gamma_{m,n}$. For $\omega \in \Delta$, let $H_\omega = \{\sigma | \sigma \in H$ and $\omega\sigma = \omega\}$, i.e., H_ω is the stabilizer of ω. We let $\bar{\Delta}$ be the subset of Δ consisting of those ω for which

$$\sum_{\sigma \in H_\omega} \chi(\sigma) = \nu(\omega) \neq 0.$$

In other words (since χ is a character of degree 1), $\bar{\Delta}$ is the set of ω for which χ is identically 1 on H_ω. If X is any n-square complex matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$, define

$$f_{H,\chi}(X) = \sum_{\omega \in \bar{\Delta}} \prod_{t=1}^{n} \lambda^{m_t(\omega)}, \tag{3}$$

the S-function of the eigenvalues defined by H and χ. (The integer $m_t(\omega)$ is the number of times t occurs in ω.) For example, if $H = S_m$ and $\chi \equiv \varepsilon$, then $\bar{\Delta} = Q_{m,n}$ is the set of strictly increasing sequences of length m chosen from $1, \ldots, n$ and $f_{H,\chi}(X)$ becomes the mth elementary symmetric function $E_m(\lambda_1, \ldots, \lambda_n)$ of the numbers $\lambda_1, \ldots, \lambda_n$. Again, if $H = S_n$, $\chi \equiv 1$, then $\Delta = \bar{\Delta} = G_{m,n}$ is the set of nondecreasing sequences of length m chosen from $1, \ldots, n$ and $f_{H,\chi}(X)$ is the mth completely symmetric function $h_m(\lambda_1, \ldots, \lambda_n)$ of the eigenvalues $\lambda_1, \ldots, \lambda_n$.

Our second result is contained in the following theorem.

THEOREM 2. *If $A_1 > 0$, $A_2 \geqslant 0$ $(A_2 \neq 0)$, and $\mu_1\mu_2 \neq 0$, then*

$$|f_{H,\chi}(\mu_1 A_1 + \mu_2 A_2)| \leqslant f_{H,\chi}(|\mu_1| A_1 + |\mu_2| A_2). \tag{4}$$

Equality can hold in (4) *if and only if $\mu_1/\mu_2 > 0$.*

As a consequence of Theorem 1 we have:

COROLLARY 1. *If A and B are hermitian, $A + B \geqslant 0$, $B \geqslant A$ $(B \neq A)$, then*

$$d_\chi^H(A) \leqslant d_\chi^H(B). \tag{5}$$

If $\chi \equiv 1$ and $A + B > 0$, then the inequality (5) *is strict.*

If $\lambda^n - p_1(X)\lambda^{n-1} + p_2(X)\lambda^{n-2} \pm \cdots + (-1)^n p_n(X)$ is the characteristic polynomial of an n-square matrix X, we immediately can conclude from Theorem 2 that:

COROLLARY 2. *If $A_1 > 0$, $A_2 \geqslant 0$ $(A_2 \neq 0)$, and $\mu_1\mu_2 \neq 0$, then*

$$|p_m(\mu_1 A_1 + \mu_2 A_2)| \leqslant p_m(|\mu_1| A_1 + |\mu_2| A_2), \qquad m = 1, \ldots, n. \tag{6}$$

Equality can hold in (6) *if and only if $\mu_1/\mu_2 > 0$.*

2. PROOFS

Preliminary to the proofs of Theorems 1 and 2 we will require the following lemma, which may be of independent interest. Denote the largest singular value of a matrix X [i.e., the largest eigenvalue of $(XX^*)^{1/2}$] by $\alpha_1(X)$.

LEMMA 1. *Assume that $A_1 > 0$ and $A_2 \geqslant 0$. If $A = \mu_1 A_1 + \mu_2 A_2$ and $S = |\mu_1| A_1 + |\mu_2| A_2$, then*

$$\alpha_1(S^{-1/2}AS^{-1/2}) \leqslant 1 \tag{7}$$

with equality if and only if either A_2 *is singular or*

$$\frac{\mu_1}{\mu_2} > 0. \tag{8}$$

Proof. Let $\lambda_1(X)$ denote the maximum eigenvalue of an arbitrary hermitian matrix X. If P is any nonsingular n-square matrix and T and B are defined by $S = PTP^*$ and $A = PBP^*$, then

$$\begin{aligned}(\alpha_1(S^{-1/2}AS^{-1/2}))^2 &= \lambda_1(S^{-1/2}AS^{-1/2}S^{-1/2}A^*S^{-1/2}) \\ &= \lambda_1(S^{-1}AS^{-1}A^*) \\ &= \lambda_1(P^{*-1}T^{-1}P^{-1}PBP^*P^{*-1}T^{-1}P^{-1}PB^*P^*) \\ &= \lambda_1(T^{-1}BT^{-1}B^*). \end{aligned} \tag{9}$$

Since $A_1 > 0$ and $A_2 \geqslant 0$, there exists a nonsingular matrix Q such that $QA_1Q^* = I_n$ and $QA_2Q^* = D$, where $d = \text{diag}(d_1, \ldots, d_n)$ and the $d_i \geqslant 0$ are the eigenvalues of $A_1{}^{-1}A_2$, $i = 1, 2, \ldots, n$. Set $P = Q^{-1}$, so that

$$\begin{aligned} T &= QSQ^* \\ &= Q(|\mu_1|A_1 + |\mu_2|A_2)Q^* \\ &= |\mu_1|I_n + |\mu_2|D. \end{aligned}$$

Similarly,

$$\begin{aligned} B &= QAQ^* \\ &= \mu_1 I_n + \mu_2 D. \end{aligned}$$

Both B and T are diagonal and hence

$$\begin{aligned} T^{-1}BT^{-1}B^* &= T^{-2}BB^* \\ &= \text{diag}\left(\frac{|\mu_1 + \mu_2 d_1|^2}{(|\mu_1| + |\mu_2|d_1)^2}, \ldots, \frac{|\mu_1 + \mu_2 d_n|^2}{(|\mu_1| + |\mu_2|d_n)^2}\right). \end{aligned}$$

By (9), $\alpha_1(S^{-1/2}AS^{-1/2})$ is the largest of the numbers

$$\frac{|\mu_1 + \mu_2 d_t|^2}{(|\mu_1| + |\mu_2|d_t)^2}, \qquad t = 1, \ldots, n. \tag{10}$$

By the triangle inequality, each of the numbers (10) is at most 1. If no d_t is 0, then $\alpha_1(S^{-1/2}AS^{-1/2}) = 1$, if and only if $\mu_1/\mu_2 > 0$. On the other hand, if any $d_t = 0$, i.e., A_2 is singular, then $\alpha_1(S^{-1/2}AS^{-1/2}) = 1$ for any μ_1 and μ_2.

We proceed to the proof of Theorem 1.

In what follows we shall use the notations, definitions, and results found in Section 3 of our paper [5, pp. 428–431] or in Section 2 of [2, pp. 164–170].

Let $V = V_n$ be the space of n-tuples of complex numbers. We use the standard inner product in V_n. Consider the linear transformations on V defined by $v \to A_i^T v$, $i = 1, 2$, i.e., the transformations on V_n whose matrix representations on the standard basis $e_t = (\delta_{1t}, \ldots, \delta_{nt})$, $t = 1, \ldots, n$ are A_i. We shall not distinguish notationally between these linear transformations and the matrices A_i since confusion is not likely to occur. Let $A = \mu_1 A_1 + \mu_2 A_2$ and $S = |\mu_1|A_1 + |\mu_2|A_2$. Now, if u is any tensor of unit length in the symmetry class $V_\chi^m(H)$, then

$$\begin{aligned} |(K(S^{-1/2}AS^{-1/2})u, u)| &\leqslant \alpha_1(K(S^{-1/2}AS^{-1/2})) \\ &\leqslant (\alpha_1(S^{-1/2}AS^{-1/2}))^m \\ &\leqslant 1, \end{aligned} \tag{11}$$

by Lemma 1. Now let $w = K(S^{1/2})e^*$, where $e^* = e_1 * \cdots * e_n$ (i.e., $m = n$), and set $u = w/\|w\|$. Then, from (11), we have

$$\begin{aligned} |(K(A)e^*, e^*)| &\leqslant \|K(S^{1/2})e^*\|^2 \\ &= (K(S^{1/2})e^*, K(S^{1/2})e^*) \\ &= (K(S^{1/2})K(S^{1/2})e^*, e^*) \\ &= (K(S)e^*, e^*). \end{aligned} \tag{12}$$

Hence

$$|d_\chi^H((Ae_i, e_j))| \leqslant d_\chi^H((Se_i, e_j)). \tag{13}$$

However, $(Ae_i, e_j) = a_{ij}$, $(Se_i, e_j) = s_{ij}$, and (13) becomes (2).

Suppose that $A_2 > 0$ and equality holds in (2). Then equality must hold in (12) which, by (11), implies that $\alpha_1(S^{-1/2}AS^{-1/2}) = 1$. Hence, by Theorem 1, $\mu_1/\mu_2 > 0$. Conversely, if $\mu_1/\mu_2 > 0$, then (2) is clearly equality for any A_1 and A_2. It remains to prove that in case $\chi \equiv 1$ the condition $\mu_1/\mu_2 > 0$ is necessary even when A_2 is singular. Assume then that A is singular and equality holds in (2). Then equality must hold in (12), so that we can write

$$\left|\left(K(S^{-1/2}AS^{-1/2})\frac{w}{\|w\|}, \frac{w}{\|w\|}\right)\right| = 1, \tag{14}$$

where $w = K(S^{1/2})e^*$. But (14) and (11) imply that $\alpha_1(K(S^{-1/2}AS^{-1/2})) = 1$. However, in general, if a linear transformation has the property that the value of the quadratic form on a vector is equal in modulus to the maximum singular value, then this vector must be an eigenvector for the transformation. Hence $w/\|w\|$ is an eigenvector of $K(S^{-1/2}AS^{-1/2})$ corresponding to some eigenvalue z of unit modulus, i.e.,

$$K(S^{-1/2}AS^{-1/2})K(S^{1/2})e^* = zK(S^{1/2})e^*,$$

$$K(A)e^* = zK(S)e^*,$$

or, equivalently,

$$Ae_1 * \cdots * Ae_n = zSe_1 * \cdots * Se_n. \tag{15}$$

It is proved in [4] that (15) implies that there exist constants d_i, $i = 1, 2, \ldots, n$, and a permutation $\sigma \in S_n$ such that $\prod_{i=1}^n |d_i| = 1$, $Se_j = d_jAe_{\sigma(j)}$, $j = 1, \ldots, n$. This means, of course, that

$$S = AQ,$$

where Q is a monomial (i.e., a product of a permutation transformation and a diagonal transformation), $|\det(Q)| = 1$. Hence

$$\begin{aligned} S &= |\mu_1|A_1 + |\mu_2|A_2 \\ &= \mu_1 A_1 Q + \mu_2 A_2 Q, \end{aligned}$$

and thus

$$\begin{aligned}\det(S) &= |\det(\mu_1 A_1 + \mu_2 A_2)|\, |\det(Q)| \\ &= |\det(\mu_1 A_1 + \mu_2 A_2)|,\end{aligned}$$

i.e.,

$$\det(|\mu_1| A_1 + |\mu_2| A_2) = |\det(\mu_1 A_1 + \mu_2 A_2)|. \tag{16}$$

Let $H = A_1^{-1/2} A_2 A_1^{-1/2} \geqslant 0$ and denote the eigenvalues of H by $h_1, h_2, \ldots, h_n$. Then (16) becomes

$$\prod_{i=1}^{n} (|\mu_1| + |\mu_2| h_i) = \prod_{i=1}^{n} |\mu_1 + \mu_2 h_i|.$$

But by the triangle inequality each term on the left is greater than or equal to the corresponding term on the right. Since $H \neq 0$, i.e., $A_2 \neq 0$, there exists a t such that $h_t > 0$. Hence $|\mu_1| + |\mu_2| h_t = |\mu_1 + \mu_2 h_t|$ implies that $\mu_1/\mu_2 > 0$.

To prove Corollary 1, observe that $B = A_1 + A_2$ and $A = A_1 - A_2$, where $A_1 = (A + B)/2$, $A_2 = (B - A)/2$. Then

$$\begin{aligned}d_\chi{}^H(B) &= d_\chi{}^H(A_1 + A_2) \\ &\geqslant d_\chi{}^H(A_1 - A_2) \\ &= d_\chi{}^H(A).\end{aligned}$$

We proceed to the proof of Theorem 2. We use the fact that, if X has eigenvalues $\lambda_1, \ldots, \lambda_n$, then

$$\operatorname{tr} K(X) = f_{H,\chi}(X).$$

This is immediate because the eigenvalues of $K(X)$ are the homogeneous products

$$\prod_{t=1}^{n} \lambda_t^{m_t(\omega)}, \qquad \omega \in \bar{\Delta}.$$

Now let $v_1, \ldots, v_n$ be an orthonormal basis of eigenvectors of S corresponding to the eigenvalues $\beta_1, \ldots, \beta_n$, respectively. By (11),

$$\alpha_1(K(S^{-1/2}AS^{-1/2})) \leqslant 1.$$

Hence

$$1 \geqslant \left| \left(K(S^{-1/2}AS^{-1/2}) \sqrt{\frac{h}{\nu(\omega)}}\, v_\omega{}^*, \sqrt{\frac{h}{\nu(\omega)}}\, v_\omega{}^* \right) \right|$$

$$= \frac{1}{\prod_{t=1}^{n} \beta_t{}^{m_t(\omega)}} \left| \left(K(A) \sqrt{\frac{h}{\nu(\omega)}}\, v_\omega{}^*, \sqrt{\frac{h}{\nu(\omega)}}\, v_\omega{}^* \right) \right|. \tag{17}$$

Thus, by the discussion immediately following (14), equality can hold for a particular ω if and only if $v_\omega{}^*$ is an eigenvector of $K(S^{-1/2}AS^{-1/2})$ corresponding to an eigenvalue of modulus 1. Thus equality holds for a particular ω, if and only if

$$K(A) \sqrt{\frac{h}{\nu(\omega)}}\, v_\omega{}^* = \xi_\omega \prod_{t=1}^{n} \beta_t{}^{m_t(\omega)} \sqrt{\frac{h}{\nu(\omega)}}\, v_\omega{}^*, \tag{18}$$

where $|\xi_\omega| = 1$.

Summing (17) for all $\omega \in \bar{\Delta}$, we have

$$\begin{aligned} f_{H,\chi}(S) &= \sum_{\omega\in\bar{\Delta}} \prod_{t=1}^{n} \beta_t{}^{m_t(\omega)} \\ &\geqslant \sum_{\omega\in\bar{\Delta}} \left| \left(K(A) \sqrt{\frac{h}{\nu(\omega)}}\, v_\omega{}^*, \sqrt{\frac{h}{\nu(\omega)}}\, v_\omega{}^* \right) \right| \\ &\geqslant \left| \sum_{\omega\in\bar{\Delta}} \left(K(A) \sqrt{\frac{h}{\nu(\omega)}}\, v_\omega{}^*, \sqrt{\frac{h}{\nu(\omega)}}\, v_\omega{}^* \right) \right| \\ &= |\operatorname{tr} K(A)| \\ &= |f_{H,\chi}(A)|. \end{aligned}$$

In the preceding calculation we have used the fact that the sum of the quadratic forms for a linear transformation over an orthonormal basis of the space is always equal to the trace. Suppose equality holds in (4). Then (18) must hold for every $\omega \in \bar{\Delta}$. Therefore

$$K(A) \sqrt{\frac{h}{\nu(\omega)}}\, v_\omega{}^* = \xi_\omega K(S) \sqrt{\frac{h}{\nu(\omega)}}\, v_\omega{}^*.$$

It follows that

$$K(A) = K(S)D,$$

where $D\sqrt{h/(\nu(\omega))}\, v_\omega{}^* = \xi_\omega \sqrt{h/(\nu(\omega))}\, v_\omega{}^*$, and hence D is a diagonal unitary transformation. Since S and D are nonsingular, we can conclude that A is nonsingular and hence has a unique polar factorization: $A = HQ$. Then

$$\begin{aligned} K(A) &= K(H)K(Q) \\ &= K(S)D, \end{aligned}$$

so that $K(Q) = D$ and hence

$$K(A) = K(SQ). \tag{19}$$

It follows [4] that Q is a monomial, i.e., a generalized permutation, and since D is diagonal we conclude that Q is diagonal also. It is easy to see that

$$\begin{aligned} I &= DD^* \\ &= K(Q)K(Q^*) \\ &= K(QQ^*) \end{aligned}$$

and thus Q is unitary. From (19) it follows that $A = \theta SQ$ for some θ, $\theta^m = 1$, i.e., for the matrices A_1 and A_2 we have

$$\mu_1 A_1 + \mu_2 A_2 = \theta(|\mu_1| A_1 + |\mu_2| A_2)Q.$$

Since $A_1 > 0$ and $A_2 \geqslant 0$, $A_2 \neq 0$, there exists an integer t such that

$$\eta(\mu_1(A_1)_{tt} + \mu_2(A_2)_{tt}) = |\mu_1|(A_1)_{tt} + |\mu_2|(A_2)_{tt},$$

where $|\eta| = 1$, $(A_1)_{tt} > 0$ and $(A_2)_{tt} > 0$. Thus

$$|\mu_1(A_1)_{tt} + \mu_2(A_2)_{tt}| = |\mu_1|(A_1)_{tt} + |\mu_2|(A_2)_{tt},$$

and it follows by the triangle inequality that $\mu_1/\mu_2 > 0$. If $\mu_1/\mu_2 > 0$, then equality clearly holds in (4).

To prove Corollary 2 we simply specialize (4) to the case $H = S_m$, $\chi = \mathrm{sgn}$.

REFERENCES

1 William M. Frank, *Proc. Amer. Math. Soc.* **16**(1965), 360–363.
2 Marvin Marcus, Lengths of tensors, in *Inequalities*, Academic Press, New York, 1967, pp. 163–176.
3 Marvin Marcus and Henryk Minc, *Trans. Amer. Math. Soc.* **116**(1965), 316–329.
4 Marvin Marcus and Henryk Minc, *J. Algebra* **5**(1967), 59–71.
5 Marvin Marcus and Henryk Minc, *Linear Algebra and Appl.* **1**(1968), 421–444.

Received September, 1970